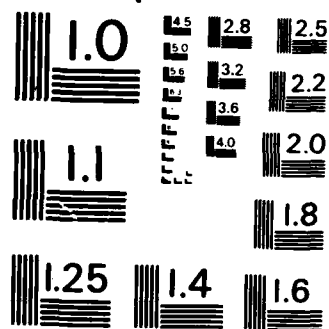END

FILMED

DTIC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

INFERENCE FOR AN EXPERIMENT BASED ON REPEATED MAJORITY VOTES

BY

JESSICA UTTS

TECHNICAL REPORT NO. 366

OCTOBER 15, 1985

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

1. INTRODUCTION

Majority vote is a decoding scheme used to enchance signal transmission in information theory. Binary signals are sent over a channel (called a binary symmetric channel) in which each signal is correctly received with probability p. To increase the probability of correct reception, each bit is repeated 2k-1 times, and the bit is decoded as the majority vote, i.e. the signal (zero or one) which is received at least k times.

If $p > 1/2$, the majority vote scheme can produce transmission accuracy which is arbitrarily good by choosing k to be large enough. Define the probability that the majority vote is correct as

$$p_m = \sum_{j=k}^{2k-1} \binom{2k-1}{j} p^j (1-p)^{2k-1-j}. \tag{1.1}$$

Then $p_m \to 1$ as $k \to \infty$, and as long as $p > 1/2$ and $k > 1$, $p_m > p$. See McEliece (1977), p. 3 for proofs and further discussion.

Puthoff (1984) recently conducted an experiment in which the majority vote technique was used to try to enchance signal transmission via ESP. Theoretically, the idea of the experiment was that for each of n repetitions, a percipient (the term parapsychologists prefer instead of "subject") would make five independent guesses as to the red/black outcome of a single spin of a roulette wheel. The response chosen at least three times would be recorded as the percipient's guess for that repetition, thus the guess would be the majority vote of five trials.

The actual experiment was slightly more complicated, for two reasons. First, if percipients were asked to guess the same target five times, they would probably repeat the same answer five times, and the concept of five independent repetitions would be completely negated. Second, the majority

vote decision could be reached in as few as three trials, so the extra guesses would be extraneous.

The actual experiment proceeded as follows. The experimenter spun the roulette wheel, immediately closing the drawer in which it was kept before it stopped spinning, so that he did not know the outcome. Thus, the experiment was designed to detect clairvoyance, extrasensory knowledge of objective events, and not telepathy. The percipient was given an HP-41C calculator which was programmed in advance to randomly assign the designations "red" and "black" to two preselected keys. The random assignment was made for each individual trial. Thus, the percipient's job was to push whichever of the two keys "felt right" for each trial. The calculator kept an internal tally of the number of times "red" and "black" had been guessed, and terminated the sequence when three of one or the other had been accumulated. The percipient then handed the calculator to the experimenter, who opened the drawer, compared the calculator result with the correct color, and told the percipient whether or not the majority vote had been successful. We will return to this example momentarily.

In general, consider an experiment which consists of n repetitions of a randomly stopped Bernoulli sequence, where each sequence is stopped at the kth success or k^{th} failure, whichever comes first. Another example of such an experiment might be one in which n poll takers each sample individuals until they get either k affirmative responses or k negative responses to some question.

The main goal in such experiments would usually be to make inferences about p, the per trial probability of success. However, there may be instances where the goal is to make inferences about p_m, the probability of correct majority vote. Notice that p_m is also the probability that an

individual sequence ends with a "success", since each sequence is terminated at the kth success or k^{th} failure, whichever comes first.

In the ESP experiment, the second goal was of interest because the whole idea of the experiment was to try to increase the overall probability of correct reception through repetition. One way to approach this goal might be to estimate p, and then use (1.1) to estimate p_m. However, with the experimental setup used, if paranormal abilities do exist they may operate on the sequence as a whole, and the trials may not be independent so that inferences about p_m would be made without the independence assumption.

The remainder of the paper is divided into four sections. In Section 2, the experiment is modeled in two different ways. In Section 3 we determine the maximum likelihood estimator for p and study its properties. Other estimators of p and p_m are derived in Section 4, and the ESP experiment is analyzed in Section 5.

2. APPROACHES TO THE PROBLEM

The experiment described above can be modeled in various ways. We consider it first as a sequence of Bernoulli trials, and then as a multinomial experiment. Although the models are theoretically the same, they are conceptually and notationally different, and it is interesting and mathematically convenient to consider both.

Let (Y_i, T_i) be the number of successes and the number of trials, respectively, needed to achieve either k successes or k failures for sequence i, i=1,...,n. Let (Y,T) be the generic random vector representing the i.i.d. random vectors (Y_i, T_i). Further, let $Y^* = \sum_{i=1}^{n} Y_i$ and $T^* = \sum_{i=1}^{n} T_i$ be the total number of successes and trials, respectively, for the entire experiment.

Define $Z_{11},\ldots,Z_{1T_1}; \ldots ; Z_{n1},\ldots,Z_{nT_n}$ as the actual randomly stopped Bernoulli sequences; then it is immediately obvious that $T_1,\ldots,T_n$ and T^* all fit the definition of stopping times. Also $Y_i = \sum_{j=1}^{T_i} Z_{ij}$; $i=1,\ldots,n$, and $Y^* = \sum_{i=1}^{n} \sum_{j=1}^{T_i} Z_{ij}$ are all randomly stopped sums. We use these facts in Section 3.

We can also model the experiment as multinomial, with each of the n sequences falling into one of 2k possible categories. The category is determined by the length of the sequence $(k,k+1,\ldots,2k-1)$, and whether it ended in success or failure. Let $\underline{X} = (X_1,X_2,\ldots,X_{2k})'$ be the 2kx1 vector of multinomial counts, where X_{2j-1} = number of sequences with exactly k failures and $(j-1)$ successes and X_{2j} = number of sequences with exactly k successes and $(j-1)$ failures, $j=1,\ldots,k$. For example, X_1 and X_2 represent the numbers of sequences composed of all failures and all successes, respectively.

Note that $\sum_{j=1}^{k} (X_{2j-1} + X_{2j}) = n$.

Define the corresponding vector of multinomial probabilities as $\underline{\pi}$, using the same subscripts. Thus, we obtain the cell probabilities:

$$\pi_{2j-1} = P(k \text{ failures, } j-1 \text{ successes, ending with failure})$$
$$= \binom{k + j-2}{j-1} p^{j-1} q^{k};$$

$$\pi_{2j} = P(k \text{ successes, } j-1 \text{ failures, ending with success})$$
$$= \binom{k + j-2}{j-1} p^{k} q^{j-1};$$

(2.1)

$j = 1,\ldots,k$, where $q = 1 - p$.

We can relate the two models by noting that

$$Y^* = \sum_{j=1}^{k} [(j-1) X_{2j-1} + kX_{2j}]$$

and   (2.2)

$$T^* = \sum_{j=1}^{k} (k + j-1) (X_{2j-1} + X_{2j}).$$

In the next few sections we will utilize both models as necessary.

3. MAXIMUM LIKELIHOOD ESTIMATOR

Since maximum likelihood estimators based on data sampled with random stopping rules are identical to those for fixed sampling, it follows from the Bernoulli model that the MLE of p is

$$\hat{p} = Y^*/T^*.$$ (3.1)

From (2.2) we obtain the relationship:

$$\hat{p} \equiv g(\underline{X}) = \sum_{j=1}^{k} [(j-1)X_{2j-1} + kX_{2j}]/ \sum_{j=1}^{k} [(k+j-1)(X_{2j-1} + X_{2j})].$$ (3.2)

3.1 Preliminary Results

We state, as lemmas, some well known results which will be used to prove the asymptotic normality of $\hat{p}$.

Lemma 3.1: If $\underline{X}$ has a multinomial distribution with n trials and probability vector $\underline{\pi}$, then $n^{-1/2} \underline{X} \xrightarrow{D}$ MVN $(\underline{\pi}, \Sigma)$ where $\Sigma = \{\sigma_{ij}\}$ is a singular matrix with $\sigma_{ij} = \pi_i(1-\pi_i)$ if $i=j$ and $\sigma_{ij} = -\pi_i\pi_j$ if $i \neq j$.

Proof: This follows from the multivariate Lindeberg-Lévy Central Limit Theorem. See e.g., Serfling (1980, p. 108).

Lemma 3.2 (Delta method): Let $\hat{Q}_n$ be an S-dimensional random vector such that $\sqrt{n}\ \hat{Q}_n \longrightarrow MVN(Q,\Sigma)$. Let $f(\hat{Q}_n)$ be a function from ^{S} to ^{1}, differentiable at $\hat{Q}_n = Q$; define $\underline{d} = [\frac{\partial f}{\partial \hat{Q}_i}|\hat{Q} = \underline{Q}]$. Then $\sqrt{n}\ (f(\hat{Q}_n) - f(\underline{Q}))$ $\longrightarrow N(0,\ \underline{d}'\ \Sigma\ \underline{d})$.

Proof: See e.g., Bishop, Fienberg and Holland (1975, p. 493).

Lemma 3.3 (Wald's equation): If Z_i, $i > 1$, are i.i.d. random variables with $E|Z| < \infty$, and T is a stopping time for Z_1, Z_2,..., then $E\left(\sum_{i=1}^{T} Z_i\right) = E(T)\ E(Z)$.

Proof: See e.g., Ross (1983, Corollary 7.2.3)

Lemma 3.4: Let W_i, $i > 1$, be i.i.d. random variables with $E(W) = 0$ and $E(W^2) = \sigma^2 < \infty$. If T is a stopping variable and S_T is the corresponding randomly stopped sum, then $E(S_T^2) = \sigma^2 E(T)$.

Proof: See Chow, Robbins and Teicher (1965, Theorem 2).

3.1 The Asymptotic Distribution of the MLE

We are now ready to prove:

Theorem 3.1: For the experiment described above, $\hat{p} = Y^*/T^*$ is asymptotically $N(p,\ pq/E(T^*))$, i.e., $\sqrt{n}\ (\hat{p}-p) \longrightarrow N(0,pq/ET)$.

Proof: Let $\hat{\underline{\pi}} = n^{-1}\underline{X}$, so (3.2) can be represented as $\hat{p} = g(\underline{X}) \equiv f(\hat{\underline{\pi}})$. Combine Lemma 3.1 and Lemma 3.2, with $\hat{Q}_n = \hat{\underline{\pi}}$, $f(\hat{Q}_n) = f(\hat{\underline{\pi}}) = \hat{p}$, $Q = \underline{\pi}$, Σ given by Lemma 3.1 and the π_i in formula (2.1), and $\underline{d} = [\frac{\partial f(\hat{\underline{\pi}})}{\partial \hat{\pi}_j}|\hat{\underline{\pi}} = \underline{\pi}]$. This results in:

$$\hat{p} \text{ is asymptotically } N(f(\underline{\pi}), n^{-1}\underline{d}' \Sigma \underline{d}). \qquad (3.3)$$

Next note that $f(\underline{\pi}) = \sum_{j=1}^{k} [(j-1)\pi_{2j-1} + k\pi_{2j}]/[\sum_{j=1}^{k} (k+j-1)(\pi_{2j-1} + \pi_{2j})]$

$= E(Y)/E(T)$. Apply Lemma 3.3 with $Z_i \sim$ Bernoulli (p), and T the stopping

time for an individual sequence, to get $E(Y) = pE(T)$, and thus $f(\underline{\pi}) = p$.

To determine the asymptotic variance in (3.3), note that $\Sigma = $ [diag. $\underline{\pi}$ -

$\underline{\pi}\underline{\pi}'$], where diag. $\underline{\pi}$ is a 2k x 2k matrix with the elements of $\underline{\pi}$ on the diagonal

and 0 elsewhere. Also, using the notation d_{2j-1} and d_{2j} to correspond to the

elements of $\underline{d}$, we find

$$d_{2j-1} = [(j-1) - p(k+j-1)]/E(T)$$

$$d_{2j} = [k-p(k+j-1)]/E(T). \qquad (3.4)$$

Thus, $\underline{d}' \Sigma \underline{d} = \underline{d}'(\text{diag } \underline{\pi}) \underline{d} - (\underline{d}'\underline{\pi})^2$

$$= \sum_{j=1}^{k} (d_{2j-1}^2 \pi_{2j-1} + d_{2j}^2 \pi_{2j}) - (\sum_{j=1}^{k} (d_{2j-1}\pi_{2j-1} + d_{2j}\pi_{2j}))^2 \qquad (3.5)$$

But notice from (3.4) that for each of the 2k multinomial outcomes, the

numerator of the corresponding element of $\underline{d}$ is simply [# successes -

p(# guesses)]. Thus, (3.5) can be rewritten as

$$[E(Y-pT)^2 - [E(Y-pT)]^2]/[E(T)]^2 = E(Y-pT)^2/[E(T)]^2 \qquad (3.6)$$

since $E(Y) = pE(T)$.

We now invoke Lemma 3.4. Let

$$W_i = \begin{cases} q \text{ if trial i is a success} \\ -p \text{ if trial i is a failure} \end{cases}$$

Then $E(W) = pq - pq = 0$, $E(W^2) = pq^2 + p^2q = pq$, and $S_T = \sum_{i=1}^{T} W_i = qY - p(T-Y)$
$= Y - pT$. Thus, from Lemma 3.4,

$$E(S_T^2) = E(Y-pT)^2 = E(W^2)E(T) = pqE(T). \qquad (3.7)$$

Combining (3.5), (3.6), and (3.7), $\underline{d}' \Sigma \underline{d} = pq/E(T)$. Since $E(T^*) = nE(T)$, we have from (3.3), $\hat{p}$ is asymptotically $N(p, pq/E(T^*))$.

3.3 Computational Formulas for E(T)

To use Theorem 3.1 for inference, we require an expression for $E(T^*)$ for fixed p and k. Although the definitional formula given by (3.11) below is computable, it is interesting to note that it can be rewritten in the computational form (3.9), via the recursion in (3.8). In particular, the closed form expression (3.10) can be used when $p = .5$, as in the ESP example. To show the dependence on k, we write ET as $E(T_k)$. The redundancy of this notation with $E(T_1)$ used earlier should not cause confusion, since the distinct roles of the subscripts 1 and k remain fixed throughout the paper.

Theorem 3.2:

$$E(T_{k+1}) = (k+1)k^{-1} E(T_k) + \binom{2k}{k} (pq)^k \qquad (3.8)$$

and

$$E(T_k) = k \sum_{j=0}^{k-1} \frac{1}{j+1} \binom{2j}{j} (pq)^j. \qquad (3.9)$$

When $p = .5$, $E(T_k) = 2k[1-\binom{2k}{k}/4^k]$. $\qquad (3.10)$

Proof: $E(T_k) = \sum_{j=1}^{k} (k+j-1) (\pi_{2j-1} + \pi_{2j})$

$$(3.11)$$

$$= k \sum_{j=0}^{k-1} \binom{k+j}{j} (p^jq^k + p^kq^j) \overset{\text{def}}{=} kS(k).$$

Now we will show that

$$(pq)^{-k}[S(k+1) - S(k)] = (k+1)^{-1} \binom{2k}{k}. \tag{3.12}$$

Using the facts that $\binom{k+j+1}{j} = \binom{k+j}{j} + \binom{k+j}{j-1}$ and $(p^{j-k}q + pq^{j-k}) = (p^{j-k} + q^{j-k}) - (p^{j-k+1} + q^{j-k+1})$, write $(pq)^{-k} S(k+1)$ in three pieces:

$$(pq)^{-k} S(k+1) = \sum_{j=0}^{k} \binom{k+j+1}{j} (p^{j-k}q + pq^{j-k})$$

$$= \sum_{j=0}^{k} \binom{k+j}{j} (p^{j-k} + q^{j-k}) \tag{3.13}$$

$$+ \sum_{j=1}^{k} \binom{k+j}{j-1} (p^{j-k} + q^{j-k}) \tag{3.14}$$

$$- \sum_{j=0}^{k} \binom{k+j+1}{j} (p^{j-k+1} + q^{j-k+1}) \tag{3.15}$$

Note that (3.13) is simply $(pq)^{-k} S(k) + 2\binom{2k}{k}$.

Also, by rewriting (3.15) so that the summation is from $j=1$ to $j=k+1$, (3.14) and (3.15) combine to give $- \binom{2k+1}{k}$. Hence, $(pq)^{-k} (S(k+1) - S(k)) = 2\binom{2k}{k} - \binom{2k+1}{k} = (k+1)^{-1} \binom{2k}{k}$, thus verifying (3.12).

From (3.12), we have $E(T_{k+1}) = (k+1)S(k+1) = (k+1)S(k) + \binom{2k}{k} (pq)^k = (k+1)k^{-1}E(T_k) + \binom{2k}{k}(pq)^k$, thus verifying (3.8).

Now use induction to show (3.9). Since, $E(T_1) = 1$, (3.9) holds for $k=1$. Suppose $E(T_k)$ can be written as in (3.9). Then from (3.8), $E(T_{k+1}) = (k+1)k^{-1} E(T_k) + \binom{2k}{k}(pq)^k$, which is precisely (3.9) for $E(T_{k+1})$.

For the special case $p = .5$ invoke the combinational identities

$$2^{-2j} \binom{2j}{j} = \Gamma(j + .5)/\sqrt{\pi} \; \Gamma(j+1); \tag{3.16}$$

see Abramowitz and Stegun (1964, p. 258, 6.1.49); and

$$\sum_{m=0}^{N} \Gamma(\beta+m)/\Gamma(\alpha+m) = (\beta-\alpha+1)^{-1} \left[\frac{\Gamma(\beta+N+1)}{\Gamma(\alpha+N)} - \frac{\Gamma(\beta)}{\Gamma(\alpha-1)} \right]; \tag{3.17}$$

see Mangulis (1965, p. 60, 24). Start with (3.9) for $p = .5$, apply (3.16), then use (3.17) with $\beta = .5$ and $\alpha = 2$, and finally use (3.16) again, resulting in the formula (3.10). //

If tables of the incomplete beta function are available, we can use those to compute ET. Note that

$$S(k) = \sum_{j=0}^{k-1} \binom{k+j}{j} \left(p^j q^k + p^k q^j \right)$$

$$= 2 - \sum_{j=k}^{\infty} \binom{k+j}{j} \left(p^j q^k + p^k q^j \right).$$

But $\sum_{j=r}^{\infty} \binom{k+j}{j} p^j q^{k+1} = I_p(r, k+1)$. Thus, $S(k) = 2 - q^{-1} I_p(k, k+1) - p^{-1} I_q(k,k+1)$ and $E(T_k) = kS(k)$.

It is interesting to note that $S(k)$ is a partial sum of the Catalan number generating function, $C(x) = \left(1 - \sqrt{1-4x}\right)/2x = \sum_{j=0}^{\infty} (j+1)^{-1} \binom{2j}{j} x^j$; see Riordan (1968). Also note that for large k, (3.10) can be approximated by

$$E(T_k) = 2k - 2(k/\pi)^{1/2}, \tag{3.18}$$

using Stirling's approximation.

3.4 Bias of the MLE

Unlike in the corresponding fixed sample size problem, the MLE $\hat{p}$ is biased. The general form of bias is $\text{Cov}(T^*, \hat{p})/ET^*$, since $\text{Cov}(T^*, \hat{p}) = \text{Cov}(T^*, Y^*/T^*)$

$= EY^* - ET^*E\hat{p} = pET^* - ET^*E\hat{p} = (ET^*)(\text{bias})$. To give some notion of the complexity of this formula, note that when $n=1$ we find

$$\text{bias} = pq(q-p)/3 \qquad \text{for } k=2$$

$$\text{bias} = pq(12pq+5)(q-p)/20 \qquad \text{for } k=3.$$

In general, bias < 0 when $p > .5$, bias > 0 when $p < .5$. When $p = .5$, bias $= 0$. This is clear since $E(\hat{p}+\hat{q}) = E[(Y^*/T^*) + (T^*-Y^*)/T^*] = 1$, but when $p = .5$ $E\hat{p} = E\hat{q}$, so $E\hat{p} = .5 = p$.

4. OTHER ESTIMATORS

One way to find an unbiased estimator is to find one for each sequence, then average over the n sequences. Note that (Y_1, T_1) is a minimal sufficient statistic for p, for sequence 1. Also, $E(Z_{11}) = p$. Thus, an unbiased estimator for p based on one sequence is

$$\hat{p}_{S1} = E(Z_{11}|Y_1, T_1) = \begin{cases} \dfrac{k-1}{T_1-1} & \text{if } Y_1 = k \\[2ex] \dfrac{Y_1}{T_1-1} & \text{if } Y_1 < k. \end{cases}$$

An unbiased estimator of p is thus

$$\hat{p}_u = \frac{1}{n} \sum_{i=1}^{n} \hat{p}_{S1}.$$

Notice that $\hat{p}_u$ is based only on what Kremers (1985) calls the "preterminal" data, i.e., the trials before the stopping trial. In particular, $\hat{p}_{S1}$ $= (\Sigma \text{ preterminal } Z_{ij})/\# \text{ preterminal trials}$.

An obvious analog of $\hat{p}_{S1}$ based on all of the preterminal trials is $\hat{p}_{pT} =$ $(Y^*-n_s)/(T^*-n) = (\# \text{ preterminal successes})/\# \text{ preterminal trials}$, where $n_s =$ number of sequences ending in success. But notice that $\hat{p}_{pT}$ is not even asymptotically unbiased. Rewriting it as

$$\hat{p}_{pT} = \frac{\sum\limits_{j=1}^{n} (j-1)X_{2j-1} + (k-1)X_{2j}}{\sum\limits_{j=1}^{n} (k+j-2)\left(X_{2j-1} + X_{2j}\right)}$$

and applying Lemmas 3.1 and 3.2, we find $E(\hat{p}_{pT}) \longrightarrow E(Y^*-n_s)/E(T^*-n)$. But $E(Y^*-n_s) = n(pET-p_m)$ and $E(T^*-n) = n(ET-1)$, where p_m is given by (1.1). Thus, asymptotically, $E(\hat{p}_{pT}) < p$ iff $p_m/p > 1$, i.e., iff $p > \frac{1}{2}$, and $E(\hat{p}_{pT}) > p$ iff $p < \frac{1}{2}$.

Finally, consider estimating p_m. In particular, suppose that we are not willing to assume that trials are independent Bernoulli, but rather that sequences are independent and fall into one of the 2k multinomial outcomes. As mentioned in the introduction, this may be a more reasonable assumption in the ESP experiment, since paranormal abilities, if they exist, may operate on the entire sequence as a whole.

With this assumption, inferences about p_m are now simply inferences about the binomial probability that the majority vote result is a success. Thus, the best estimator for p_m is simply $\hat{p}_m = \frac{1}{n} \sum\limits_{j=1}^{k} X_{2j}$ and its variance is $p_m(1-p_m)/n$.

5. ESP EXPERIMENT

The experiment described in the introduction was carried out with k = 3 and n = 100, for three percipients. Results are shown in Table 1. We would like to test the hypothesis that p = .5. From (3.10), $ET^* = 412.5$. Also, $Var(\hat{p}_u) = .027$, derived by writing out its distribution for p = .5, k = 3, n = 100. Table 1 also gives values of the test statistic Z based on the three estimators $\hat{p}$, $\hat{p}_u$, $\hat{p}_m$. Notice that in each case $\hat{p}_u < \hat{p}$, which would be expected to be the trend if $p > \frac{1}{2}$, since $\hat{p}$ is biased.

It is also interesting to see if the independence assumption is reasonable for the data. Table 2 presents the results for two sets of chi-square tests. For the first set, H_0:Guesses are independent, $p = .5$ was tested, and for the second set $\hat{p}$ was used in place of .5, so that just the independence hypothesis was being tested. Notice that the results indicate that the independence assumption seems reasonable for all three percipients, but the assumption that $p = .5$ is questionable for two of them. The Z tests also support this conclusion.

One curious result is that percipient #2 had about the expected number of sequences of length 5, but the split between ending in success and ending in failure was 27 to 12 instead of about half of each. This led to $Z = 2.00$ for the test of $p_m = .5$, whereas both tests of $p = .5$ resulted in nonsignificance. The parapsychologists used this as evidence that a test based on p_m would be more powerful than one based on p. This is clearly not true if the independence assumption is valid, since the test based on p_m results in loss of information.

Table 1: Results for ESP Experiment

Outcome			Percipient	
S	F	1	2	3
0	3	4	12	9
3	0	15	11	17
1	3	13	16	22
3	1	32	22	26
2	3	12	12	9
3	2	24	27	17
Y^*, T^*		250,417	219,416	220,400
$\hat{p}, Z$		.600,4.06	.526,1.06	.550,2.03
$\hat{p}_u, Z$		.587,3.21	.505,.185	.547,1.73
$\hat{p}_m, Z$		.710,4.20	.600,2.00	.600,2.00

Table 2: Chi-square Tests of Independence Assumption

a. H_0:Trials are independent, p = .5

Outcome			P1		P2		P3	
S	F	Exp.	Obs.	X^2	Obs.	X^2	Obs.	X^2
0	3	12.5	4	5.78	12	.02	9	.98
3	0	12.5	15	.5	11	.18	17	1.62
1	3	18.8	13	1.79	16	.42	22	.55
3	1	18.8	32	9.27	22	.55	26	2.76
2	3	18.8	12	2.46	12	2.46	9	5.11
3	2	18.8	24	1.44	27	3.58	17	.17
				21.24		7.21		11.19

d.f. = 5   p < .005   p > .10   p ≅ .05

b. H_0:Trials are independent

Outcome			P1			P2			P3	
S	F	Exp.	Obs.	X^2	Exp.	Obs.	X^2	Exp.	Obs.	X^2
0	3	6.4	4	.9	10.6	12	.18	9.1	9	.00
3	0	21.6	15	2.02	14.6	11	.89	16.6	17	.01
1	3	11.5	13	.20	16.8	16	.04	15.0	22	3.27
3	1	25.9	32	1.44	20.7	22	.08	22.5	26	.54
2	3	13.8	12	.23	17.7	12	1.84	16.5	9	3.41
3	2	20.7	24	.53	19.6	27	2.79	20.2	17	.51
				5.32			5.82			7.74

d.f. = 4   p ≅ .25   p ≅ .20   p ≅ .10

REFERENCES

Abramowitz, M. and Stegun, I. (1964) Handbook of Mathematical Functions,
National Bureau of Standards, Applied Math Series #55.

Bishop, Y.M.M., Fienberg, S. E., and Holland, P. W. (1975)   Discrete
Multivariate Analysis, MIT Press, Cambridge, MA.

Chow, Y.S., Robbins, H., and Teicher, H. (1965)   "Moments of Randomly Stopped
Sums," Annals of Mathematical Statistics, 36, 789-799.

Kremers, W. K. (1985) "The Statistical Analysis of Sum-Quota Sampling,"
Ph.D. Thesis, Cornell University.

Mangulis, V. (1965)   Handbook of Series for Scientists and Engineers, Academic
Press, New York.

McEliece, R. J. (1977) The Theory of Information and Coding, Addison-Wesley,
Reading, MA.

Puthoff, H. E. (1984) Calculator-Assisted Psi Amplification, Technical Report,
SRI International, Menlo Park, CA.

Riordan, J. (1968) Combinatorial Identities, John Wiley and Sons, New York.

Ross, S. M. (1983) Stochastic Processes, John Wiley and Sons, New York.

Serfling, R. J. (1980) Approximation Theorems of Mathematical Statistics,
John Wiley and Sons, New York.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 366	2. GOVT ACCESSION NO. AD-A161 141	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Inference For An Experiment Based On Repeated Majority Votes		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Jessica Utts		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0475
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 411SP		12. REPORT DATE October 15, 1985
		13. NUMBER OF PAGES 18
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, If different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ESP Experiment, stopping time, maximum likelihood estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consider an experiment which consists of n independent Bernoulli sequences, each of which is randomly terminated at either the kth success or the kth failure, whichever comes first. The goals are to make inferences about the per trial probability of success, as well as the probability that a sequence is terminated with a success. The latter is equivalent to the probability that if $2k-1$ trials are accumulated, the majority will be successes. Various estimators are investigated, and the results of an ESP experiment fitting this scheme are examined.

DD 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

12-85

DTIC